# Accounting for uncertainty and correlation in earthquake loss estimation

P. Bazzurro
*AIR Worldwide Co., San Francisco, California, USA*

Nicolas Luco
*AIR Worldwide Co., San Francisco, California, USA (now at US Geological Survey, Golden, Colorado, USA)*

*Keywords: Earthquake loss estimation, treatment of uncertainty, correlation, portfolio losses, seismic risk mitigation and management*

ABSTRACT: Quantifying the potential impact of earthquakes on portfolios of properties located in seismically prone regions is of primary interest to property owners, insurance and reinsurance companies, capital lending institutions, local government agencies, and structural engineers, among others. Each is likely to have a different viewpoint and different requirements. Owners and corporate risk managers can cope with seismic risk using a variety of strategies. These range from establishing self-insurance programs (e.g., expense funds for post-earthquake emergency or proactive seismic retrofit), to buying earthquake insurance coverage and, more recently, to taking advantage of alternative risk transfer arrangements, including captives and insurance-linked securities, or catastrophe bonds. Regardless of which risk transfer mechanism is ultimately chosen, it is critical that the estimates of potential loss on which these decisions are based are as accurate as possible given the available information. Often earthquake loss analyses are conducted using numerical approaches that balance speed of execution with rigorousness of the methodology. This balance can sometimes cause undesirable portfolio-specific effects on the accuracy of the loss estimates. A rigorous treatment of uncertainty is often the aspect that suffers the most. This article presents two examples of the effects on the loss estimates of considering or neglecting the different sources of uncertainty and correlation. In particular, we consider ground motion uncertainty for a given set of earthquake parameters (e.g., magnitude and source-to-site distance), ground motion correlation at different sites during the same event, uncertainty in structural response given the level of ground motion, uncertainty in loss given a level of structural response, and correlation between losses to similar buildings in the same portfolio. The effects on commonly used loss metrics, such as Average Annual Loss (AAL) and loss Exceedance Probability (EP) curves, are discussed.

## 1 INTRODUCTION

Until the 1980's, loss estimates to portfolios of properties caused by natural catastrophes such as earthquakes, hurricanes, tornadoes, and flood were extrapolated from those caused by historical events. However, the limited span covered by historical catalogs, the lack of systematically gathered loss data, and the rapidly increasing building inventory in regions of high risk led to severe underestimation of such losses. One such case is the 1992 Hurricane Andrew that hit Florida and unexpectedly caused 25 to 30B USD in economic losses and drove several insurance companies into dire straits. As a result, purely actuarial approaches for the estimation of losses generated by rare natural catastrophes were abandoned in favor of probabilistic engineering tools now widely applied.

Today it is not uncommon for primary insurance companies to perform loss estimation analyses of large portfolios that contain as many as a million properties. At the other end of the spectrum corporate risk managers may deal with a single building or a handful of buildings located at different sites. In both portfolio cases multiple properties may be affected by the same event, a consideration that makes the loss computation more complex. In this article, we consider only cases where the location of each property and the building type are known. In many applications, however, mainly involving reinsurance company portfolios, these basic characteristics are often not available. In these cases, the assumptions that the analyst is forced to make (e.g., locating the properties

at the population-weighted centroid of an often very large geographical area) are so pervasive that the accuracy of the loss estimates is difficult to assess.

Even when location and building type are known, the analyses are often performed without proper consideration for the many sources of uncertainty and correlation that are present in the problem. In this article we will look into some of these issues. Although many aspects inherent to loss estimations are conceptually identical for different natural perils, we limit ourselves here to discussing losses generated by earthquakes. We will also only discuss estimates of the ground-up (i.e., total) losses before any insurance deductibles or limits are applied.

## 2 SINGLE-SITE VS. PORTFOLIO LOSS ESTIMATION PROCEDURES

### 2.1 *Single structure*

Before embarking on a discussion of portfolio loss estimates, we first summarize an approach that can be adopted for assessing the earthquake losses of a single structure at a given site. Step 1 involves modeling the future earthquakes that may affect the structure at the site along with their relative annual frequency of occurrence. This step requires knowledge about the location of faults, past historical seismicity, and rate of recurrence of earthquakes of different magnitude. Step 2 evaluates the distribution of the possible ground motion at the site generated by each earthquake. This step is based on the knowledge of both the basic characteristics of the earthquake, such as magnitude, $M$, and source-to-site distance, $R$, and the soil conditions at the site. Step 3 involves estimating the response of the structure for any given level of ground motion intensity, as measured by an appropriate parameter such as spectral acceleration. The structural response is gauged by quantities that are known to correlate well with damage, such as maximum interstory drift and maximum absolute floor acceleration. Step 4 assesses the monetary losses (and sometimes other types of consequences, such as casualties and downtime) that can result from the level of structural and non-structural damage caused by any given level of building deformation.

In practical applications each step can be carried out at different levels of detail (e.g., ASTM, 1999) that are commensurate with the available analysis budget. Regardless of the level of sophistication of the analysis, at least four sources of uncertainty can be present: First, the uncertainty in the parameters of the earthquake recurrence model and fault locations. Second, the ground motion variability at the site for any given $M$-$R$ scenario (Figure 1). If the analysis assumes a generic soil at the site, this uncertainty is described by the dispersion measure in a ground motion attenuation relationship. Third, the response record-to-record variability of a structure subject to different ground motions of the same intensity (e.g., same spectral acceleration) (Figure 2). Fourth, the uncertainty in the losses for a structure that has experienced a certain level of deformation (e.g., a maximum interstory drift of 1%) (Figure 3).

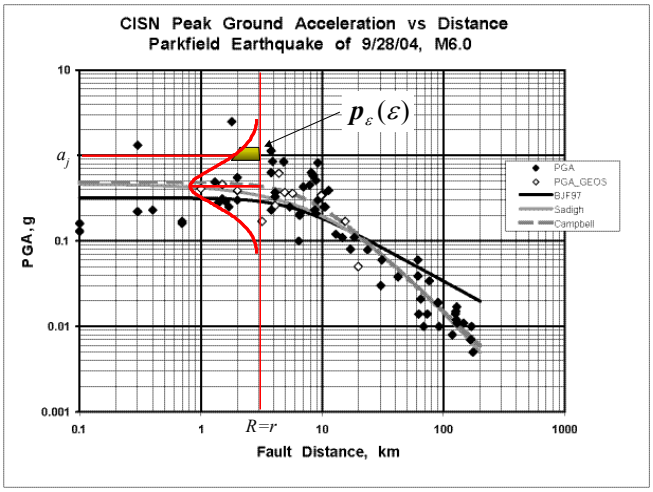


**Figure 1:** Variability of a ground motion parameter, here Peak Ground Acceleration (PGA), on similar soil conditions generated by the recent M6.0 September 28, 2004 Parkfield Earthquake (data courtesy of Tony Shakal of the California Geological Survey). The three lines provide the median PGA predicted by three widely used attenuation relationships for shallow crustal events. The standard normal variable $\varepsilon$ captures the variability of the ground motion conditional on $M$ and $R$ (in this example, M6.0 and 3.0km, respectively).

The objective of the probabilistic loss estimation is computing a relationship between the monetary loss, $L$, and the annual Mean Rate of Exceedance (MRE) or, alternatively, the Mean Return Period (MRP). The MRP is defined as the reciprocal of the annual MRE (e.g., MRE of $4 \times 10^{-4}$/year implies a MRP of 2,500 years). The first source of uncertainty listed above is epistemic in nature and is usually dealt with via a logic tree approach (e.g., SSHAC, 1997). The loss computations are performed multiple times according to the parameter values defined by each combination of the branches of the logic tree and the final results are pooled according to the weights of each branch combination. Here, we will consider this type of uncertainty only for Step 1 (i.e., the earthquake recurrence model and fault locations).

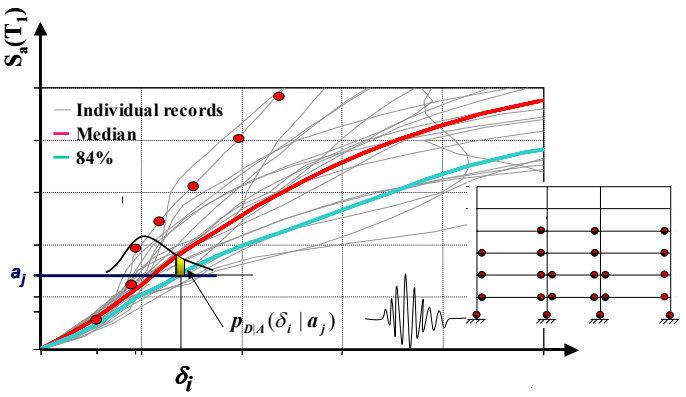


**Figure 2:** Example of variability of the structural response measure, here the maximum interstory drift, $D$, caused by different records having the same value of the spectral acceleration, $S_a$, at the fundamental frequency of vibration, $T_1$. Each thin line in the figure is produced by shaking the structure with the same ground motion record scaled by increasingly higher factors, and each time recording the value of $S_a$ and $D$ (see red dots on one of the lines). The pattern of damage in the building frame for one such shaking analysis is also shown in the figure. (Part of this figure is courtesy of Prof. Krawinkler of Stanford Univ.)

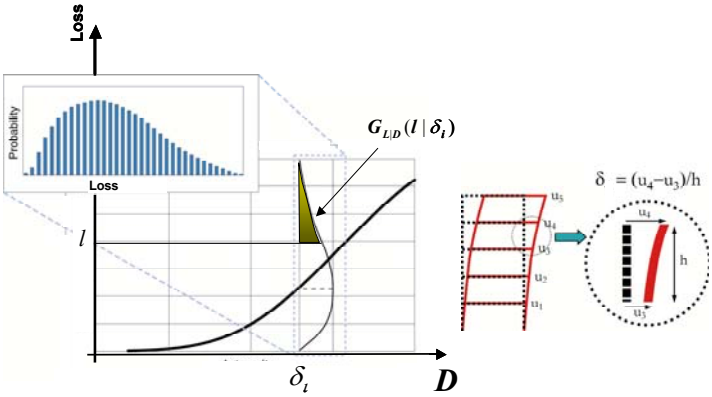


**Figure 3:** Sketch of a loss function that probabilistically relates the severity of the response, as measured by the parameter $D$ (e.g., the maximum peak interstory drift along the height of the building), with the loss suffered by the structure. The figure on the right shows the interstory drift at the 4th story of a building.

For each branch combination, the other three uncertainty sources considered here are integrated according to an equation that, after removing some dependencies that in most cases have been (empirically) shown to be either non-existent or mild, can be written in discrete form as:

$$G_L(l) = \sum_{all\,\delta_i} \sum_{all\,a_j} G_{L|D}(l \mid \delta_i) p_{D|A}(\delta_i \mid a_j) p_A(a_j) \qquad (1)$$

$\boldsymbol{G_{L|D}(l \mid \delta_i)}$ is the complementary cumulative distribution function (CCDF) of loss, $L$, given the level of structural response, measured here by the maximum (along the height of the building) peak (over time) interstory drift, $D$ (see Figure 3). It is assumed that $L$ is conditionally independent of the ground motion parameter, $A$. This CCDF can be estimated by pricing the appropriate repair strategies for the level of damage in any structural (e.g., columns) and non-structural (e.g., partitions) element in the building that is expected for the level of structural deformation $\delta_i$. For example, for small values of $\delta_i$ partitions may develop only minor cracks that can be fixed by simply patching and repainting.

$\boldsymbol{p_{D|A}(\delta_i \mid a_j)}$ is the probability that $D$ is "equal to" (or, more precisely, "in the neighborhood of") the value $\delta_i$ given that the ground motion parameter, $A$, at the site is "equal to" $a_j$ (see Figure 2). In this paper we consider $A$ to be the spectral acceleration $S_a$ at the fundamental structural period $T_1$. To numerically evaluate this conditional probability one can utilize multiple ground motion records increasingly scaled to higher values of $S_a(T_1)$, as shown in Figure 2. Empirical studies (e.g., Shome *et al.*, 1998) have shown that for many structures $D \mid S_a(T_1)$ is conditionally independent of (or very mildly dependent on) $M$ and $R$. This implies that different records with the same $S_a(T_1)$ value but generated by different earthquakes (i.e., different $M$-$R$ pairs) induce, on average, the same value of $D$ in most structures. Hence, the potential dependence on $M$ and $R$ has been dropped.

$\boldsymbol{p_A(a_j)}$ is the annual probability that the maximum value of $S_a(T_1)$ at the site is "equal to" $a_j$. This term can be obtained by numerically differentiating the conventional seismic hazard curve for $S_a(T_1)$ (and taking the absolute value of the result), which can be computed for the site using any Probabilistic Seismic Hazard Analysis (PSHA) code. Alternatively, this term can be evaluated directly by adding the hazard from $n$ seismic sources that generate events with different $M$-$R$ pairs, according to the following equation:

$$\boldsymbol{p_A(a_j) = \sum_{k=1}^{n} \nu_k \left( \sum_{all\,m} \sum_{all\,r} \sum_{all\,\varepsilon} I_A(a_j \mid m, r, \varepsilon) p_{M,R}(m, r) p_\varepsilon(\varepsilon) \right)_k} \qquad (2)$$

The quantity $\nu_k$ is the mean annual rate of occurrence of earthquakes generated by source $k$ with magnitude greater than some specific lower bound (e.g., $M$5.0). $\boldsymbol{I_A(a_j \mid m, r, \varepsilon)}$ is an "indicator" function equal to 1 if $S_a$ is "equal to" $a_j$ for $M$=$m$, $R$=$r$, and $\varepsilon$=$\varepsilon$, according to the attenuation relationship considered for the fault $k$, and equal to 0 otherwise. Recall that an attenuation equation can be written as $\ln(S_a) = g(M, R, \theta) + \varepsilon \sigma_{\ln(S_a)}$, where $\boldsymbol{g}(\cdot)$ is the functional form used when regressing the ground motion database, $\theta$ represents additional variables such as fault type, $\sigma_{\ln(S_a)}$ is the regression standard error, and

$\varepsilon$ is a standard Gaussian variable that models the scatter of $S_a \mid M, R, \theta$ (see Figure 1). $\boldsymbol{p}_{M,R}(\boldsymbol{m},\boldsymbol{r})$ is the joint probability mass function (PMF) of $M$ and $R$, which can be numerically computed by running a PSHA code for a value of $S_a$=0. $\boldsymbol{p}_{\varepsilon}(\varepsilon)$ is the PMF of $\varepsilon$, which is stochastically independent of $M$ and $R$.

Finally, note that for the earthquakes of interest the "rare event" assumption holds (i.e., the likelihood of two or more earthquakes in the unit time is small compared to the likelihood of one event only). This assumption implies that the MRE of a loss, $l$, is numerically equal to the probability that the annual maximum loss exceeds $l$. Here we make this assumption and we are therefore able to use the notation of $G_L(l)$ for this CCDF, understanding that it is numerically identical to the loss MRE curve.

The first case study in the results section of this paper considers the case of a single building to illustrate these concepts.

## 2.2 *Portfolio of structures*

The procedure for loss estimation of portfolios of properties adopts many of the concepts described above for a single site, but with three key differences.

The *first and most important difference* stems from the fact that multiple sites are, in general, affected by the same earthquake with ground motion intensities that are correlated. This makes the loss computation via Equations 1 and 2 theoretically possible but highly impractical. As in many other cases when more elegant solutions cannot be easily pursued, Monte Carlo simulation becomes the approach of choice. The *M*-*R* domain of future events, instead of being exhaustively searched as done in the PSHA approach, is sampled to generate a more limited number of events than used in a single-site loss estimation study. This limitation is due only to computational restrictions that may arise when the portfolio contains many thousands of sites.

This simulation procedure generates a catalog that contains a realization of the possible future events that may occur in a *fixed period* of time, say 10,000 years. The occurrence of future events is often modeled as a Poissonian process. For each event the rupture area is simulated according to empirical relationships (e.g., Wells and Coppersmith, 1994), and randomly positioned along the fault. If the analysis accounts for directivity effects, the location of the hypocenter is also simulated using distributions based on empirical data for each fault mechanism.

The *second key difference* is that the simulation of the stochastic ground motion field generated by an earthquake is not as well established. The ground motions at different sites are correlated, in general, in an anisotropic fashion due to (among other things) the direction of the rupture propagation. This is evident in Figure 5, which shows seismograms recorded during the 1994 *M*6.7 Northridge earthquake, where the rupture propagated upward and northward.

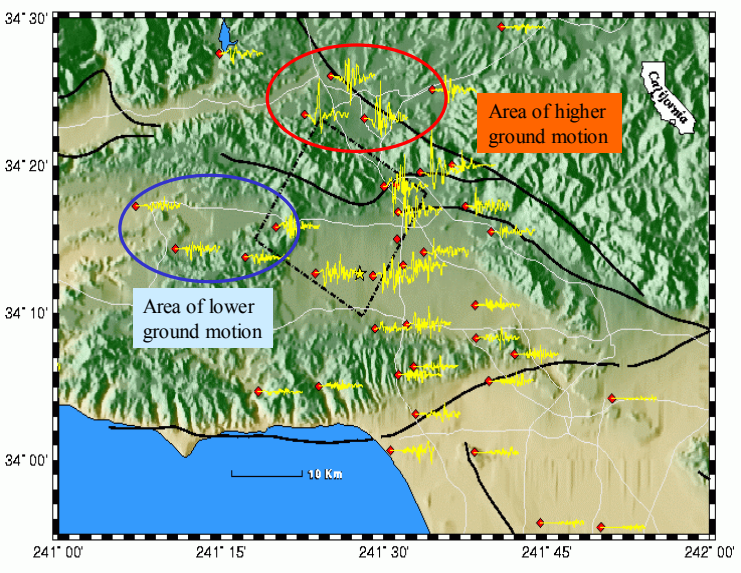


**Figure 4:** Correlated ground motions from the 1994 *M*6.7 Northridge earthquake in Southern California. The two highlighted regions, which are at approximately the same distance from the fault rupture (i.e., the rectangle in the middle of the figure), experienced dramatically different ground motions.

Anisotropic models for simulating correlated ground motion fields have been proposed in the literature (e.g., McGuire, 1988) but they are seldom used in practice due to the difficulty in estimating their parameters. More commonly utilized is the isotropic approach based on the random effects model (Abrahamson and Youngs, 1992), where the error term, $\varepsilon$, is split into two factors such that Ln $(S_a)$ = $f(M,R,\boldsymbol{\theta}) + \xi\sigma_1 + \eta\sigma_2$, where $\xi$ and $\eta$ are standard Gaussian random variables that reflect the intra-event variability (i.e., the variability within sites equidistant from the rupture) and the inter-event variability (i.e., the systematic differences in the ground motion at all sites from different events), respectively. The quantities $\sigma_1$ and $\sigma_2$ are the standard errors from the statistical regression of empirical data (e.g., Lee *et al.*, 2000). For a given attenuation relation and a given $(M,R,\boldsymbol{\theta})$ triplet, $\sigma_1$ and $\sigma_2$ are known quantities. Hence, ground motion realizations at $k$ sites from the same earthquake are obtained by drawing $k$ random samples of $\xi$ but only one sample of $\eta$.

The *third key difference* lies in the loss estimation module. In portfolio analyses buildings that are "similar" in their basic characteristics, such as construction type (e.g., steel moment-resisting frames), the number of stories, and the vintage are grouped into a so-called construction class. The building design code enforced at the time of construction implicitly captures the "average" seismic resilience to earthquakes of different buildings in the same construction class. The enforcement of the same code introduces an obvious source of correlation between losses at different buildings within the same class that is naturally accounted for by using the same loss function (see Figure 3). There are other factors unrelated to the building code that may increase the building-to-building loss correlation. For example, residential buildings in the same housing development project are likely built by the same construction company at the same time using the same crews and applying the same set of rules. In this case, the structures may *all* under- or over-perform "average" buildings in the same class. The quantification of the building-to-building loss correlation requires data that, unfortunately, are not yet available. Hence, in real applications the correlation structure is often subject to a sensitivity analysis. This source of correlation, of course, is more important for more homogeneous portfolios consisting of buildings clustered together at a limited number of sites than for portfolio of buildings spread over a large number of locations.

We conclude this section by describing how the earthquake event catalog is utilized for estimating losses to portfolios of properties. The Monte Carlo technique is adopted to generate one (correlated) ground motion random field for each earthquake in the event catalog. For each site in the portfolio the loss to each property is simulated considering building-to-building loss correlation, if appropriate. The portfolio loss induced by each earthquake is equal to the sum of the losses at all the sites. When this is done for all the earthquakes in the catalog, the portfolio losses are ranked. The highest loss can be considered as one possible realization of the portfolio loss that has a MRP of 10,000 years, or the loss that has 1 x $10^{-4}$ mean rate of being exceeded each year. The second highest loss has a mean annual exceedance rate of 2 x $10^{-4}$, and so on. Hence, a realization of the exceedance probability (EP) curve for the portfolio losses can easily be generated. The same exercise can be repeated multiple times and an alternative set of portfolio losses can be created for each simulation. Also, alternative earthquake catalogs may be generated using the same or a different set of assumptions and the loss estimation analysis repeated again. Hence, this procedure can lead to the estimation of a family of EP curves from which important statistics such as the median and the 10% and 90% percentile curves can be empirically estimated.

## 3 RESULTS

### 3.1 *Case Study 1: Single Building*

The first case study involves a large building in San Francisco, California (Figure 5). The building is a steel moment-resisting frame structure with four stories above ground and one below grade. The building footprint is approximately 75m × 160m, its height is about 36m above street level, and the total building area is 36,000m$^2$. The building total replacement cost, $C_{tot}$, is 200M USD. This building houses several commercial enterprises. The owners were interested in estimating the potential ground-up losses and corresponding likelihood of exceedance that the building may sustain as a result of earthquakes. The estimates of loss and business interruption downtime were requested to help the management make more informed decisions regarding the selection of the best risk mitigation strategy.

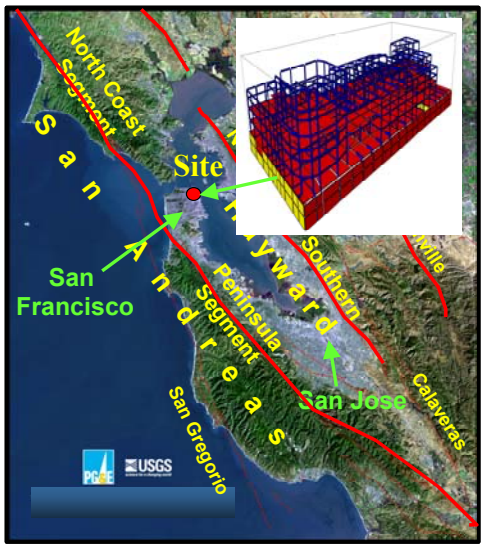


**Figure 5:** Location and 3-dimensional computer model of the structure considered in Case Study 1.

The results of the seismic hazard analysis for the building site performed using the attenuation relationship for generic soil conditions developed by Abrahamson and Silva (1997) are shown in Figure 6. The solid red line refers to the analysis performed excluding epistemic uncertainty and the ground motion

variability, namely by setting $p_{\varepsilon}(\varepsilon)=1$ for $\varepsilon=0$ and equal to zero otherwise in Equation 2 and Figure 1. All the other hazard curves are obtained by including the sources of uncertainty that were labeled above as No.1 (i.e., the uncertainty in the parameters of the earthquake recurrence model and fault locations) and No. 2 (the ground motion variability at the site for any given *M*-*R* scenario). Including the ground motion variability drastically increases the hazard at the site. The random variables *M* and *R* are bound by the physical properties and locations of the faults and they cannot freely increase. Forcing the third random variable $\varepsilon$ to be zero artificially disallows any large but realistic ground motion (in this case larger than about 0.6g) from ever happening at the site. The *M*6.0 Parkfield earthquake, for example, generated a PGA in excess of 2.5g at a site at about 2km from the fault (Figure 1), a value seven times larger than that predicted using $\varepsilon=0$. This has an obvious significant impact on the risk, as will be seen later. Performing a logic tree analysis to account for the epistemic uncertainty (uncertainty No. 1) allows us to evaluate a family of hazard curves from which we can empirically estimate the desired percentiles. Note that if we had used only the best estimates of each parameter we would have obtained a hazard curve in the central part of the bundle that would *not* have coincided with either the mean or the median curve.

The results in Figure 6 are shown in terms of exceedance rates of given ground motion levels. The EP loss curves in Figure 7 were obtained by convolving the mean $p_A(a_j)$ from Figure 6 (after differentiation) with the response and the loss parts according to Equation 1.

The curve in dark blue is "deterministic" and demonstrates that, unless the various sources of uncertainty are modeled, the most severe loss will not exceed about 15% of the building replacement value. Modeling only the ground motion variability (the values of $\sigma_{\ln(S_a)}$ are in the range of 0.6 to 0.7) and neglecting the other uncertainty sources (No. 3 and 4) results in the green curve, which clearly shows that losses significantly greater than 15% can in fact occur. In addition to the ground motion variability, the curve in red includes the variability in the response given the ground motion level. The coefficients of variation (COV's) of $D|S_a(T_1)$ increase with $S_a(T_1)$ from about 0.1 at very small drifts to about 0.6 at $S_a(T_1)$ values of about 1.0g, and slightly decrease thereafter. This is why the red and the green curves diverge at a high loss of about 35% caused by large ground motions. The curve in light blue further includes the uncertainty in the loss estimates given the structural response. The COV's of $L|D$ used in the analyses decrease from about 0.8 for small $D$ values to about 0.2 for large values of $D$ (such as 6% to 8%, when damage varying from major to collapse is expected in the building). The small COV's of $L|D$ for large $D$ values cause the light blue curve to approach the red curve for loss ratios above 60%. When all the sources of uncertainty are modeled, the collapse of the building (i.e., total loss) is estimated to occur, on average, once every 4,000 years.

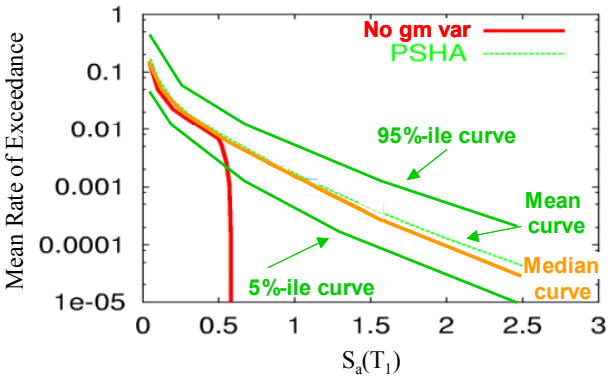


**Figure 6:** Seismic hazard curves for the building site expressed in terms of the spectral acceleration $S_a(T_1)$ at the fundamental period of vibration of the building in its longitudinal direction, which is the weakest one.

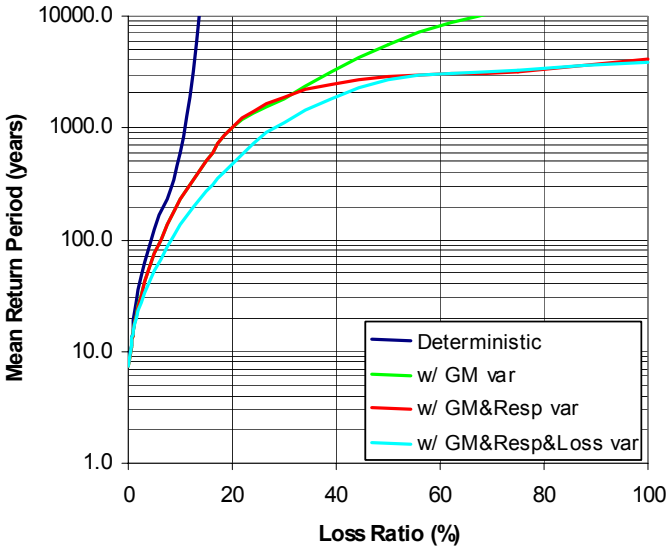


**Figure 7:** Mean Return Period of losses exceeding a given percentage of the building replacement cost. The effect of not modeling the ground motion variability (i.e., deterministic curve), as is often the case in practice, results in a severe underestimation of the risk. Legend: GM=ground motion.

The effect of the different treatments of uncertainty on the Average Annual Loss (AAL) is summarized in

Table 1. The AAL value almost doubles from the estimate of the deterministic analysis to the estimate of the analysis that accounts for all the other uncertainty sources considered here. Note that the AAL values in the second and third rows in Table 1 are very close because the AAL value is controlled by the small losses that occur more frequently.

Based on the results shown here, a risk-neutral decision maker who embraces the axioms of classical decision-making theory would decide to insure the building only if the annual premium is AAL x $C_{tot}$ or less (assuming no deductible or limit). Armed with the results of the most comprehensive analysis considered here, he/she would pay as much as 0.37 x \$200M = \$740,000 every year for earthquake insurance. On the other hand, the results from the cheaper but less detailed deterministic analysis would give him/her a false sense of security, with the consequence that the building would only be insured if the annual insurance premium is less than \$440,000.

Since most of the decision makers in real life are risk averse to some degree, the differences in the optimal decision that would stem from using the results of the different analyses would be even more striking. When the decision maker is not risk-neutral, the optimal solution is found by maximizing the decision maker's utility, rather than by minimizing the expected loss. The optimal decision for a risk-averse decision maker is driven by the fear of large losses, the likelihoods of which are grossly underestimated (see Figure 7) by the deterministic analysis and, to a lesser degree, also by the analysis that includes only the variability in the ground motion at the site. With the knowledge gained from the most refined analysis results, the risk-averse decision maker would be willing to spend significantly more than \$740,000 per year in earthquake premiums.

A retrofitting option that involves adding a few X braces was designed for this building and estimated to cost about \$6M, which is 3% of the building replacement cost. A loss estimation analysis of the retrofitted design with all of the uncertainty sources considered provided an AAL of 0.21%. If implemented, this retrofit option would drastically reduce the chances of large losses and, therefore, lower the value of earthquake insurance. For the risk-neutral decision maker the amount spent in retrofitting would be recuperated through reduced insurance premiums (i.e., 0.37%-0.21% times \$200M) in about 20 years, which is a longer time horizon than usually considered acceptable. However, in the case of an adverse risk attitude, the proactive retrofitting coupled with lower insurance premiums would more likely be the preferred choice.

| Type of Analysis | AAL (%) |
|---|---|
| Deterministic | 0.22 |
| w/ gm var | 0.31 |
| w/ gm&resp var | 0.31 |
| w/ gm&resp&loss var | 0.37 |

**Table 1:** Average Annual Loss (AAL) for the building.

### 3.2 *Case Study 2: Multiple Buildings at the Same Site*

The second case study deals with assessing the earthquake loss potential of an industrial plant in Bogotá, Colombia, with the aim of establishing the optimal earthquake risk mitigation strategy. The plant comprises a total of 25 buildings of several construction types and vintages, all located within 1km$^2$. The total replacement value of all the buildings is about 140M USD. Four types of analyses were performed with the results displayed in Figure 8: "deterministic" (black line), with ground motion variability given $M$ and $R$ (blue line), with the above plus loss variability given response measure and response variability given the ground motion level (green line), and with all of the above plus building-to-building loss correlation (red line). In this case the loss was predicted directly from the ground motion parameter without the intermediate step provided by the structural response. Hence, the conditional uncertainty of $L|A$ combined the uncertainty of $L|D$ and of D|$A$. Given their proximity, we assumed the same ground motion at each building (i.e., perfect positive correlation). The curves in Figure 8 represent the median curves from 100 simulations that here are all based on the same earthquake event set. Using additional event sets might affect the median curves and would undoubtedly increase the scatter in the EP curves.

As expected, the deterministic analysis again greatly underestimates the earthquake risk. Unlike the single-site case, however, the loss uncertainty has a larger impact on the results. This occurs because the available information about each building here does not go beyond a generic description of the basic characteristics, such as construction material, number of stories, and year built. The lack of detailed information is reflected in larger conditional uncertainty in the loss estimates given the ground motion level. The COV's of $L|A$ used in the analyses range from 0.1 at large ground motions to 4 for small ground motions. The extremely large COV values, which are applied only at very small ground motions, are based on loss data from past earthquakes. Such large COV values are due to the expected loss being very small, because

of the large percentages of buildings that did not suffer any losses. Modeling the building-to-building loss correlation also increased the chance of large losses. The increase is not very significant, however, because the correlation in this case was assumed to be rather low.

Note that modeling both the large uncertainty in the loss estimates given the structural response (which dominates this study) and the loss correlation between similar buildings has an effect on the lower end of the EP curves as well. The losses corresponding to shorter MRP values are smaller than those estimated by including only the ground motion variability in the estimation problem. Excluding the dominant uncertainty source and the loss correlation would prevent the model from predicting the occurrence of both very low and very high portfolio losses that are experienced when the losses are lower than average and higher than average at many buildings.

Interestingly, the median AAL estimates are almost identical (i.e., 0.13%) for all three probabilistic analyses. This is due to the fact that the three EP curves cross at a mean return period of about 100 years (see inset of Figure 8). The AAL from the deterministic analysis is, however, substantially lower—only 0.08%. The effects that the results from these different types of analyses may have on the optimal risk mitigation strategy are similar to those mentioned for the single-site case study.

## 4 FINAL REMARKS

This article has presented (i) a summary of a methodology for earthquake-generated loss estimation of both single structures and portfolios of structures located at different sites, and (ii) two applications, one for a single site and one for a portfolio. The focus of the paper was on the effects of modeling the different sources of uncertainty and correlation on both the annual loss exceedance probability curves and average annual losses. We have considered four different uncertainty sources and, for the portfolio case, two sources of correlation. Multiple analyses were performed between the two extremes of "deterministic" and fully probabilistic, where all uncertainty and correlation sources were sequentially considered.

As expected, the deterministic analyses severely underestimated the losses predicted by the fully probabilistic case. Furthermore, the exceedance probability curves were more affected than the AAL estimates by the modeled uncertainty sources. This is to be expected given that the AAL values are dominated by frequent small losses. Very large losses on the other hand are only observable in fully probabilistic analyses. Hence, the results of a loss estimation analysis that does not properly incorporate all the uncertainty sources may be misleading and cause the decision maker, particularly if risk averse, to make ill-advised decisions. Such decisions would be based on an underestimation of the actual risk. For a building owner, the consequences may be underinsurance and decreased motivation for proactive building strengthening. For an insurance company, the consequences may be an under-pricing of insurance premiums.

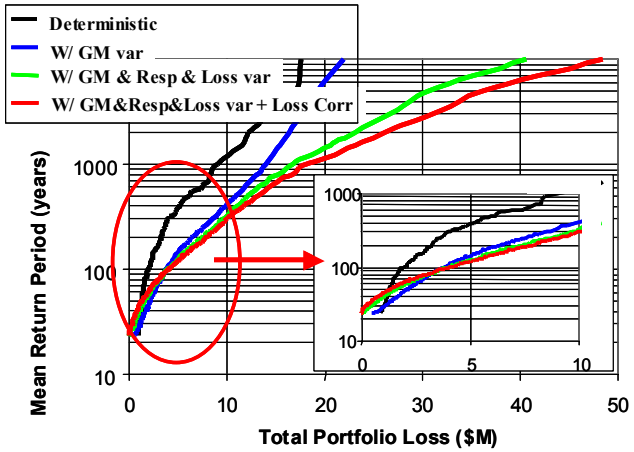


**Figure 8:** Mean Return Period of losses exceeding a given amount. All the curves other than the deterministic one are the median curves obtained form 100 simulations. Legend: GM=ground motion.

## 5 REFERENCES


Abrahamson, N. A., and W. J. Silva (1997), "Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes", *SRL.*, Vol. 68 (1), pp. 58-73.

Abrahamson, N. A., and R.R. Youngs (1992), "A Stable Algorithm for Regression Analyses Using the Random Effects Model", *BSSA*, Vol. 82, No. 1, pp. 505-510.

America Society of Testing and Materials (ASTM) (1999). *Standard Guide for the Estimation of Building Damageability in Earthquakes*, ASTM E2026-99, ASTM.

McGuire, R.K. (1988). "Seismic Risk to Lifeline Systems: Critical Variables and Sensitivities", *Proc. of 9$^{th}$WCEE*, Tokyo-Kyoto, Japan, August 2-9, Vol. VII, pp. 129-134.

Senior Seismic Hazard Analysis Committee (SSHAC) (1997). "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts", NUREG-CR-6372, UCRL-ID-122160, Vol. 1, Prepared for LLNL.

Shome, N., Bazzurro, P., Cornell, C.A., and J.E. Carballo (1998). ``Earthquakes, Records, and Nonlinear MDOF Responses", *Earthquake Spectra*, Vol.14, No.3, Aug., 469-500.

Wells, D. and K. Coppersmith (1994)."New Empirical Relationships Among Magnitude, Rupture Length, Rupture Width, and Surface Displacement", *BSSA*, Vol. 84, pp. 974-1002.